How to run AsterixDB locally on your machine

1. Preparation

You need to check the Java Development Kit (JDK) installation on your machine. You must upgrade your JDK to the latest version (**1.8**) to run AsterixDB. To do the required check:

For Windows users:

a) Start the Command Prompt (How).

b) Execute 'java -version'.

c) If you see 'java version xxxx' that means you have Java properly installed.

d) If you see 'java is not recognized as xxxx', that means you don't have Java installed. Check this to install Java and try again.

For OS X users:

a) Start your terminal.

b) Execute 'java -version'.

c) If you see 'java version xxxx' that means you have Java properly installed.

d) If you see 'command not found java', that means you don't have Java installed. Check this to install Java and try again.

For Linux users:

a) Start your terminal.

b) Execute 'java -version' to make sure Java is properly installed.

c) If you didn't install the Java environment, follow the link to install it.

d) Execute 'echo $JAVA_HOME'.

e) If you see a blank line, that means your environment variable is not set properly. Execute the following command to set it:

```
export JAVA_HOME=PATH_TO_YOUR_JAVA_INSTALLATION
```

f) You can put this command into your configuration file (e.g., .bashrc) to set the path variable automatically every time you login.

g) Rerun the command in b) and d) to make sure everything is properly configured.

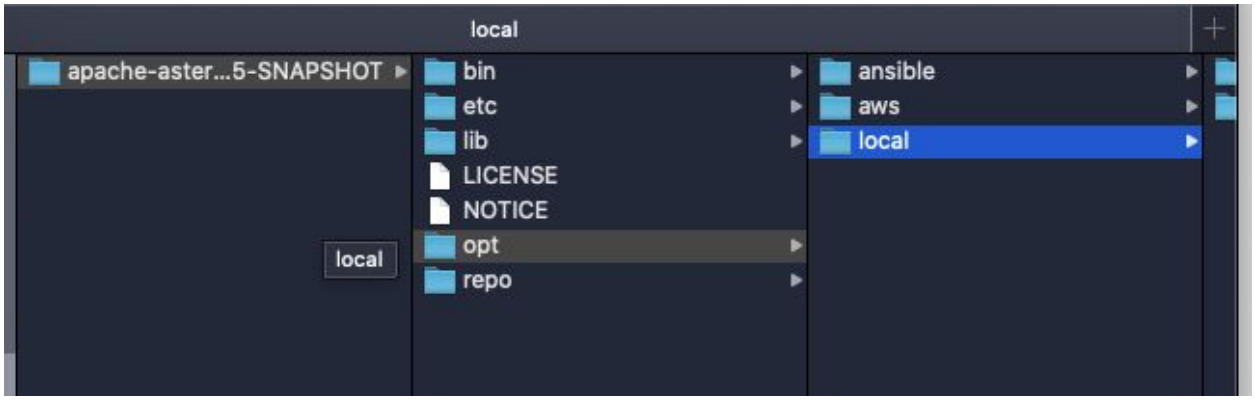
openJDK is also OK, but we highly recommend the Oracle JDK.

2. Download an AsterixDB sample cluster

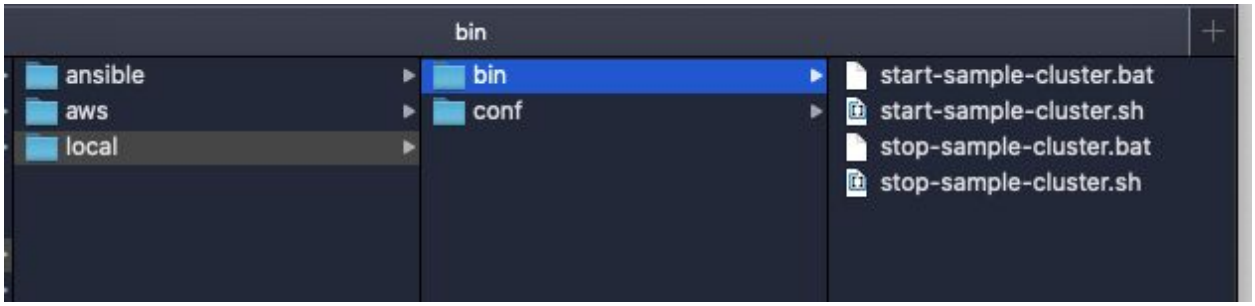
Download AsterixDB (version 0.9.5) from the link here: [Link]. Do not download from other sources!

3. Start an AsterixDB sample cluster

After you download the package, move it to your preferred working directory and unzip it. Please check the download folder if you can't find it. You should see the following structure once you're done unzipping the package:



The scripts used for starting a sample AsterixDB cluster reside in the `opt/local/bin` directory. You should see 4 scripts files under this (`opt/local/bin`) directory:



For Windows Users:

a) [**Recommended**] Start the Windows Command Prompt and navigate to `asterix-server-0.9.5-SNAPSHOT/opt/local/bin` under your working directory. Execute the command: `start-sample-cluster.bat`

For example, if you downloaded this file and uncompressed it in the Downloads folder, you may want to execute the following command. Here, XXX is your username.

```
cd c:\Users\XXX\Downloads\asterix-server-0.9.5-SNAPSHOT\opt\local\bin

start-sample-cluster.bat
```

b) You can also navigate to this directory using your file browser and double click on `start-sample-cluster.bat` to start the cluster.

c) By executing the stop-sample-cluster script in the same way, you can shutdown the sample cluster. Execute the command: `stop-sample-cluster.bat`.

For Linux/OSX Users:

a) Start your terminal.

b) Navigate to `asterix-server-0.9.5-SNAPSHOT/opt/local/bin` under your working directory.

c) Execute start-sample-cluster.sh using the following command:

```
./start-sample-cluster.sh
```

Whichever way you start the sample cluster, if everything works properly, you should see the following message printed.

```
INFO: Starting sample cluster... INFO: Waiting up to 30 seconds for cluster
127.0.0.1:19002 to be available.
INFO: Cluster started and is ACTIVE.
```

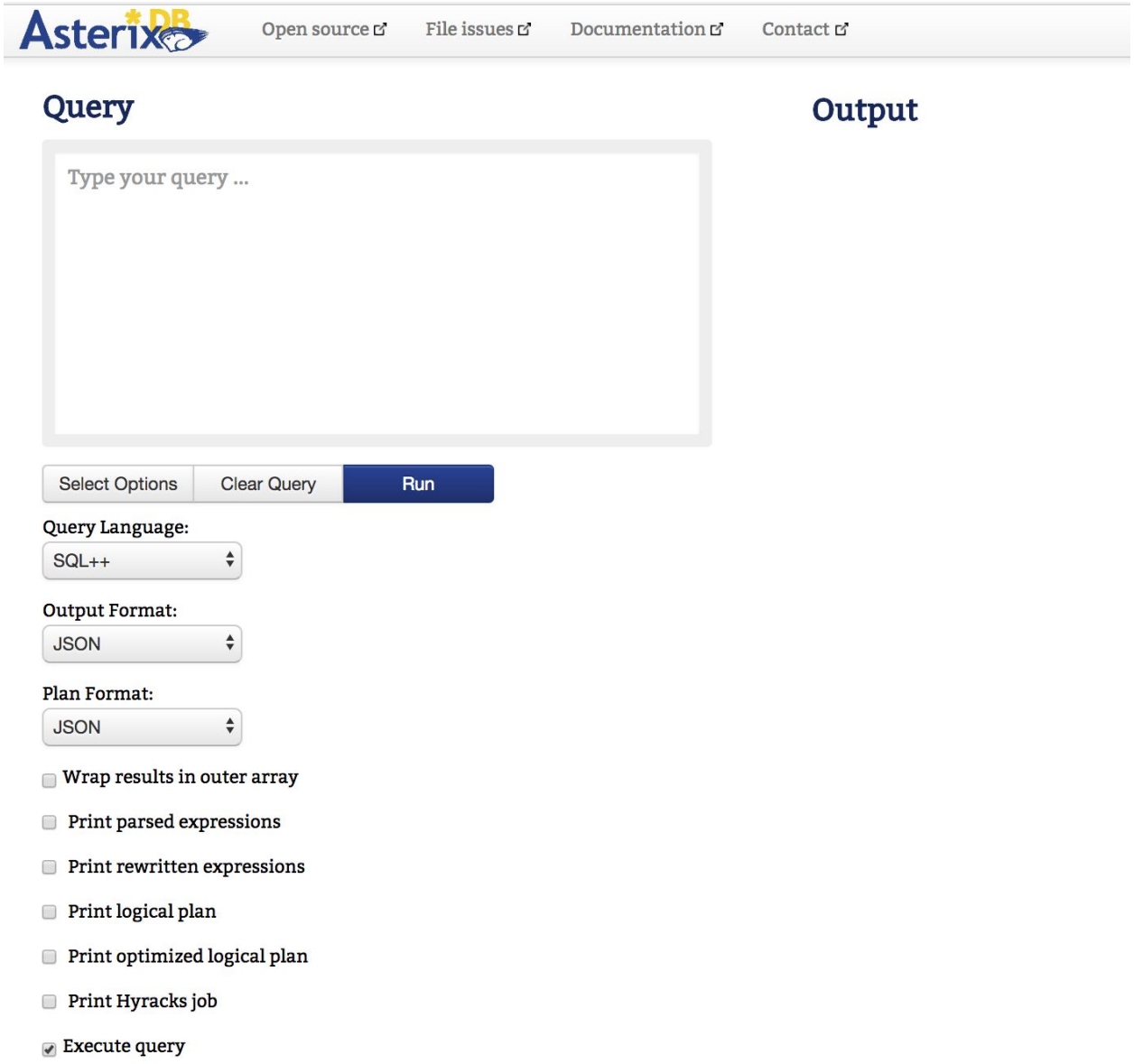
By executing the stop-sample-cluster script in the same way, you can shutdown the sample cluster.

```
./stop-sample-cluster.sh
```

Note that AsterixDB is designed to run on parallel hardware; the sample cluster is a virtual cluster with two nodes.

4. Access the AsterixDB query interface

After you have started the local cluster, you can access the query interface from your browser at: http://localhost:19001. Its query interface looks like this:



Type the following query into the query box on the left and click the `Run` button to execute it:

```
SELECT * FROM Metadata.`Dataset`;
```

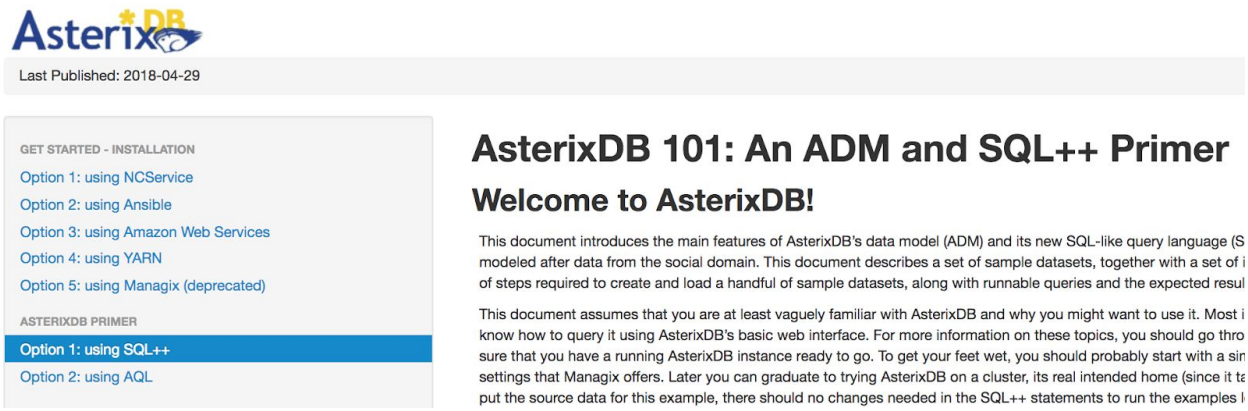
You will see the query result in the blank area on the right side. The default query language for AsterixDB is SQL++.

5. Enjoy the ride!

If you are here in the instructions, now will you have a mini-cluster running an AsterixDB instance on your machine. There are a lot more interesting features and examples to explore the Apache AsterixDB website. **NOTE: You must read and execute all of the materials covered in The SQL++ Primer before starting the actual homework**. Do not attempt the homework or ask questions about SQL++ on Piazza before you have done this step! You may also wish to look at The SQL++ Query Language and Builtin Functions while doing the homework problems. We have also attached the dataset in the SQL++ Primer to the class wiki page as a script using Insert Statements. Please give them a try and enjoy!

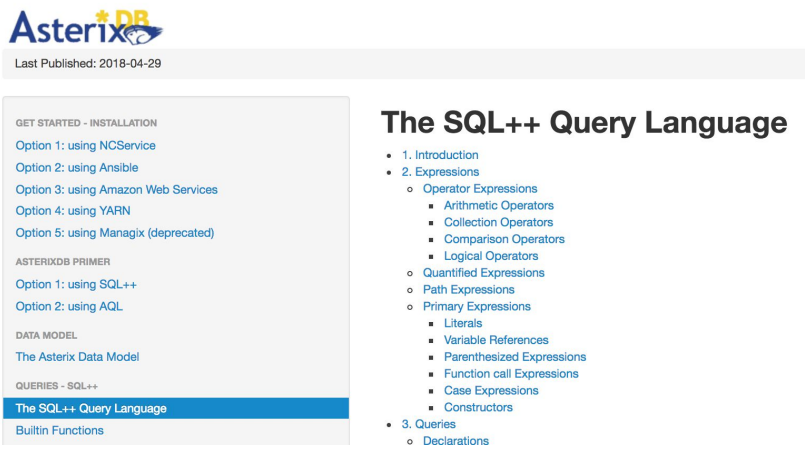
1) AsterixDB 101: SQL++ Primer

https://ci.apache.org/projects/asterixdb/sqlpp/primer-sqlpp.html



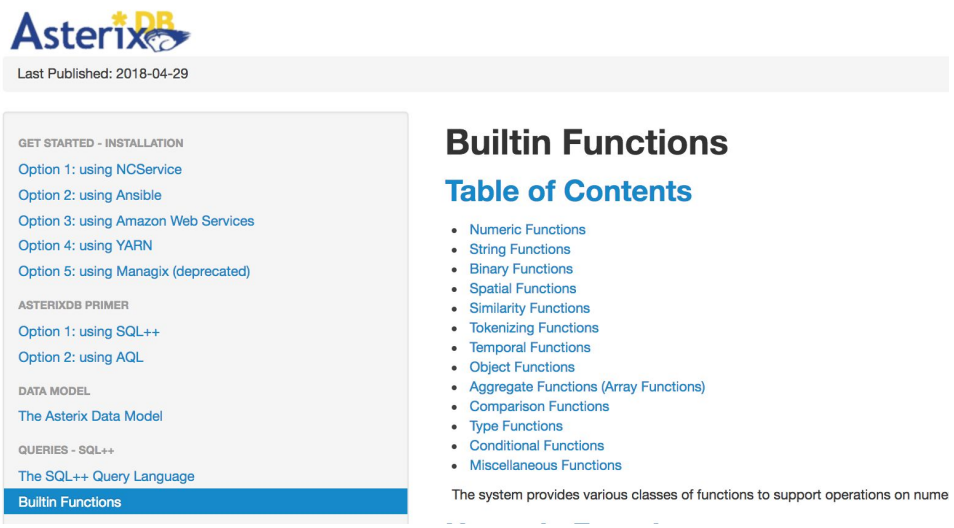
2) The SQL++ Query Language

https://ci.apache.org/projects/asterixdb/sqlpp/manual.html



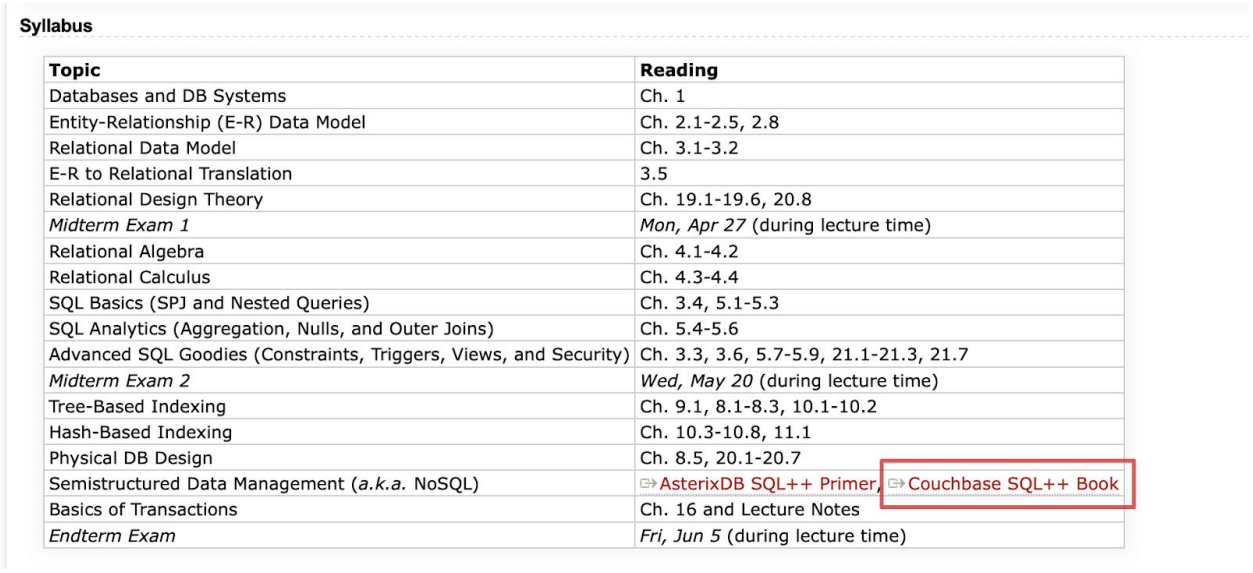
3) Builtin Functions

https://ci.apache.org/projects/asterixdb/sqlpp/builtins.html



3) SQL++ for SQL Users (Don Chamberlin)

https://asterixdb.apache.org/files/SQL_Book.pdf



Syllabus

Topic	Reading
Databases and DB Systems	Ch. 1
Entity-Relationship (E-R) Data Model	Ch. 2.1-2.5, 2.8
Relational Data Model	Ch. 3.1-3.2
E-R to Relational Translation	3.5
Relational Design Theory	Ch. 19.1-19.6, 20.8
Midterm Exam 1	*Mon, Apr 27* (during lecture time)
Relational Algebra	Ch. 4.1-4.2
Relational Calculus	Ch. 4.3-4.4
SQL Basics (SPJ and Nested Queries)	Ch. 3.4, 5.1-5.3
SQL Analytics (Aggregation, Nulls, and Outer Joins)	Ch. 5.4-5.6
Advanced SQL Goodies (Constraints, Triggers, Views, and Security)	Ch. 3.3, 3.6, 5.7-5.9, 21.1-21.3, 21.7
Midterm Exam 2	*Wed, May 20* (during lecture time)
Tree-Based Indexing	Ch. 9.1, 8.1-8.3, 10.1-10.2
Hash-Based Indexing	Ch. 10.3-10.8, 11.1
Physical DB Design	Ch. 8.5, 20.1-20.7
Semistructured Data Management (*a.k.a.* NoSQL)	AsterixDB SQL++ Primer, Couchbase SQL++ Book
Basics of Transactions	Ch. 16 and Lecture Notes
Endterm Exam	*Fri, Jun 5* (during lecture time)